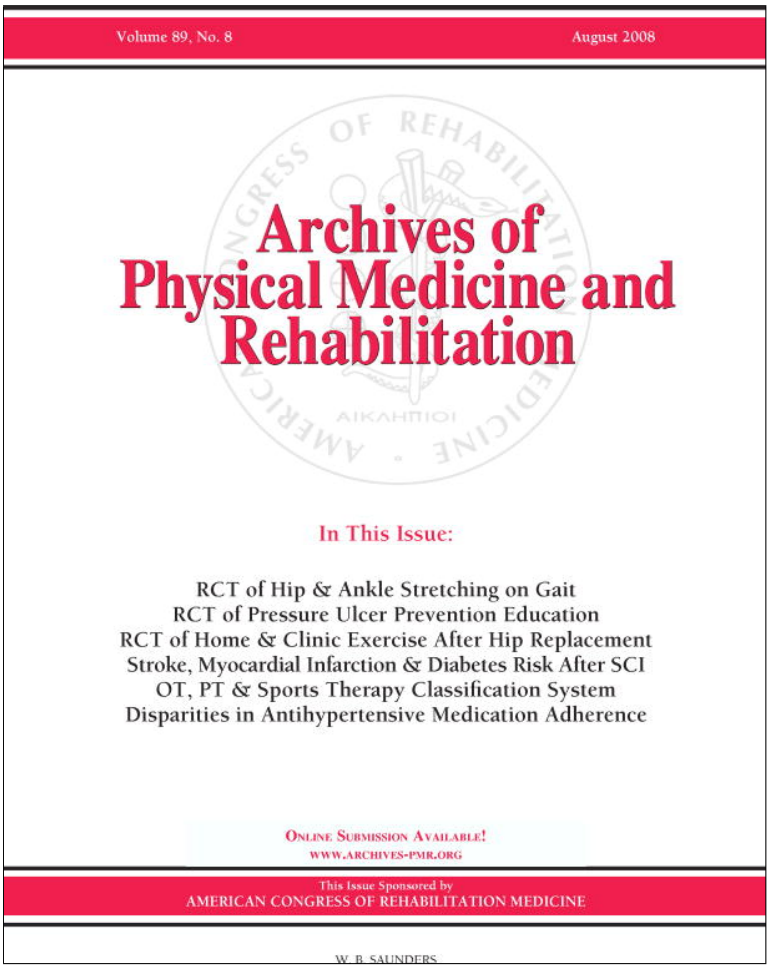

DEPARTMENTS

Letters to the Editor

The Steep Ramp Test: Questions About Sensitivity and Reliability

It was with great interest that I read the important work of De Backer et al[1] on the steep ramp test as an exercise training tool for cancer survivors. However, I have 2 questions after reading the study.

First, De Backer only describes the correlation between the steep ramp test and the submaximal exercise test and the maximal oxygen uptake ($\dot{V}O_2$max) test, before and after exercise training, and reports moderate to high correlations between the tests. However, it seems more interesting to correlate the individual training improvements as measured with the 3 tests to see whether the steep ramp test and the submaximal exercise test are tracking the improvements in $\dot{V}O_2$max of the patients (the criterion standard).

Second, I question the use of the intraclass correlation coefficient (ICC) for reliability of the steep ramp test. As indicated by Atkinson and Nevill,[2] the ICC gives high values in a heterogeneous sample. Other statistics, such as the limits of agreement or coefficient of variation, give further insight into the reliability of an instrument. Can the authors provide additional reliability statistics for the steep ramp test using the current data set?

Tim Takken, MSc, PhD
Wilhelmina Children's Hospital
University Medical Center Utrecht
Utrecht, The Netherlands
t.takken@umcutrecht.nl

Disclosure: No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

References

1. De Backer IC, Schep G, Hoogeveen A, Vreugdenhil G, Kester AD, van Breda E. Exercise testing and training in a cancer rehabilitation program: the advantage of the steep ramp test. Arch Phys Med Rehabil 2007;88:610-6.
2. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217-38.

doi:10.1016/j.apmr.2008.05.002

Flawed Analyses in Driving Pedal Reaction Time Study

Meikle et al[1] reported an interesting study on driving pedal reaction times after right transtibial amputations. Unfortunately, the data analyses are seriously flawed.

There were 10 subjects, 4 treatment conditions, and 18 test trials within each condition. The main analyses are described as three "1-way repeated measures ANOVAs [analyses of variance]."[1(p390)] This suggests that there was no real interest in any variation across the 18 trials within each condition. That is typical for response time measures, and one usually computes a median (or trimmed mean) of the raw response times within each treatment (T) by subject (S) combination. Had Meikle done so, the data matrix would consist of 10 rows (1 per subject) and 4 columns (1 median response time per condition), and the degrees of freedom (*df*) for a 1-factor repeated-measures ANOVA would be as follows[2]:

$$\text{total } df = 4 \times 10 - 1 = 39$$

$$\text{between-subject } df = 10 - 1 = 9$$

$$\text{within-subject } df = 39 - 9 = 30$$

$$\text{treatment } df = 4 - 1 = 3$$

$$\text{error } df = \text{treatment} \times \text{subject } df = 3 \times 9 = 27$$

Thus, the F tests for 1-way repeated-measures ANOVAs on these data would have 3 and 27 degrees of freedom. But the reported F tests have denominator degrees of freedom ranging from 157 to 163. How can this be?

Notice that Meikle reported Wilks λ values along with the F tests. This indicates that the author group actually performed multivariate analysis of variance (MANOVA) rather than repeated-measures ANOVA. (Textbooks often recommend MANOVA in place of repeated-measures ANOVA, because it does not require the assumption of sphericity, which is frequently violated.[2]) But this does not solve the problem of the mysterious degrees of freedom, because with 10 subjects and 4 outcome variables, multivariate F tests would have 3 and 7 degrees of freedom.

The only way to arrive at the reported degrees of freedom, as far as I can tell, is to perform MANOVA (not repeated-measures ANOVA) on a data matrix that consists of 4 columns and 180 rows (ie, 18 rows per subject), minus rows that contain outliers in any of the 4 columns. (Meikle reports that, on average, 3.5% of the data points were excluded for each outcome measure.) But of course, doing so treats the multiple rows of data for a given subject as if they are from multiple subjects who are independent of each other. This is *inappropriate*, and *seriously* violates the assumption of independence of observations between subjects.

Thus, the reported F and *P* values are not valid. The good news for Meikle is that the pattern of means will not likely be affected much by a proper analysis of the data.

Bruce Weaver, MSc
Northern Ontario School of Medicine
Lakehead University
Thunder Bay, ON, Canada
bweaver@lakeheadu.ca

Disclosure: No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the author or on any organization with which the author is associated.

References

1. Meikle B, Devlin M, Pauley T. Driving pedal reaction times after right transtibial amputations. Arch Phys Med Rehabil 2006;87:390-4.
2. Howell DC. Statistical methods for psychology. 6th ed. Belmont: Duxbury Pr; 2007.

doi:10.1016/j.apmr.2008.05.004